25
28
22
20
28
32
36
25
22
20
100 mm
1.0
1.1
1.25
1.4
1.6
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.8
ABCDEFGHIJKLMNOPQRSTUVWXYZ
abcdefghijklmnopqrstuvwxyz
1234567890
1.0 mm
1.5 mm
2.0 mm
A5
20
22
25
28
32
36
40
20
22
25
28
32
36
40

DOCUMENT RESUME

ED 314 435 TM 014 223

AUTHOR Rakow, Ernest A.; McSparrin, Brenda
TITLE Internal Consistency among Scores in the Tennessee Career Ladder Program as Measured by Effect Size.
PUB DATE Nov 89
NOTE 18p.; Paper presented at the Annual Meeting of the Mid-South Educational Research Association (Little Rock, AR, November 8-10, 1989).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Career Ladders; *Classification; Classroom Observation Techniques; *Effect Size; Elementary Secondary Education; Peer Evaluation; Principals; *Reliability; State Programs; Student Evaluation of Teacher Performance; *Teacher Evaluation; Vertical Organization
IDENTIFIERS *Internal Consistency; *Tennessee Career Ladder Program

ABSTRACT

The purpose of this paper is to discuss a procedure for determining the degree of relationship of the many scores in the Tennessee Career Ladder Program to the final classification based upon this system. The program, now in its sixth year of implementation, places over 40,000 teachers and educational administrators within a career ladder hierarchy. Sources of information pertinent to the evaluation include teacher, principal, student, and peer evaluations. Data from data collection instruments, including classroom observation and interviews, account for 75% of a candidate's total score, while a consensus reached by evaluators accounts for the remaining 25%. The sample for this analysis of the program consists of all general education teachers evaluated for upper level positions on the ladder during the 1988-89 academic year. Sixty-six scores are calculated in this system. It appears that effect size can be used to indicate the degree to which scores in a complex personnel evaluation system differentiate among the levels earned. The use of a modified level in which that data were excluded from the classification scheme reduced the magnitude of the effect sizes while generally keeping different scores in approximately the same relative order. Four data tables and three bar graphs are provided. (TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED314435

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

ERNEST A. RAKOW

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

INTERNAL CONSISTENCY AMONG SCORES IN THE TENNESSEE CAREER LADDER PROGRAM AS MEASURED BY EFFECT SIZE

Ernest A. Rakow and Brenda McSparrin

Department of Foundations of Education
Memphis State University 38152

Paper presented at the Annual Meeting of the Mid-South Educational Research Association, Little Rock, AR., November 1989

TM014223

BEST COPY AVAILABLE

INTERNAL CONSISTENCY AMONG SCORES IN THE TENNESSEE CAREER LADDER PROGRAM AS MEASURED BY EFFECT SIZE

Ernest A. Rakow and Brenda McSparrin
Memphis State University

The purpose of this paper is to discuss a procedure for determining the degree of relationship of the many scores in a complex personnel evaluation system to the final classification based upon this system. The statistic utilized is effect size, as originally developed for the purpose of meta-evaluation.

INTRODUCTION

The State of Tennessee is now in the sixth year of the implementation of the Career Ladder Program for teachers and administrators. Presently, over 40,000 educators are participating in this program. Of this number, over 8,000 have attained the top rungs of the ladder.

The placement of teacher on the various levels of the ladder is dependent on the results of performance evaluations. For the licensing levels of Probationary, Apprentice, and Career Level I the local school system is responsible for evaluating and certifying performance. Assessment for Career Levels II and III, the upper levels, is conducted by a three-member team of peers from outside the candidate's own school system.

The summative decision to be made by the state evaluation is to classify the teachers into one of three categories: Career Level III - the distinguished teacher; Career Level II - the outstanding teacher; and retention of current status, i.e. Career Level I - the competent teacher. This decision is based on a composite scale score derived through complex evaluation and scoring procedures. The domains of competence which provide the logical basis for this evaluation process are given in Table 1.

The Evaluation Process and Instrumentation

The Tennessee Career Ladder evaluation system is designed to collect information from several sources of data. These sources include the teacher, the principal, the students, and the peer evaluators. This approach adds complexity to the administration and data management aspects of the evaluation in such areas as evaluator scheduling, instrument scoring, record keeping, and analysis.

Instruments used to collect data from several sources include 1) a classroom observation record which is administered at least six times, 2) three dialogues between the evaluators and the

teacher focusing on planning, teaching strategies, and evaluation of students and instruction, 3) a summary of professional growth and leadership activities (PDL), 4) written tests of communication and professional knowledge in the various domains of competence, 5) a student questionnaire, 6) a principal questionnaire, and 7) a consensus judgment rating of competence in each domain, except leadership, by the evaluation team. The several instruments are administered during three day-long visits (one visit by each evaluator) over the evaluation cycle. These instruments are listed in Table 1.

The observation record is a mix of low inference and high inference measures. Some of the items are marked as overtly present (1) or absent (0). Other items consider qualitative aspects based upon patterns of behaviors which occurred over the observation period. These items use a metric based on a 1-5 point scale or are derived through formulas. Raw score values from the observation instrument have lower limits or 0 or 1 and upper limits of 100. The dialogues, student questionnaires, and the PDL summaries use a 1-5 point scale. The professional skills tests consists of 25 items in each of four domains of competence being measured. The principal questionnaire and the consensus rating use a 10-point scale corresponding to gradations within the various career ladder levels and range from 200-800.

The Scoring Process

The scores in the Tennessee Career Ladder Evaluation system result from a compilation of the data collected by the instruments and from a consensus judgement of the evaluators. The data from the instruments account for 75 percent of a candidates's total score while the consensus accounts for the remaining 25 percent. In order to determine whether a teacher qualifies for upper career levels, two things must be accomplished. The teacher's score by domain of competence must be calculated and translated to a 200-800 reporting scale and it must be determined if a composite score derived from all the domains is high enough to qualify for Career Level II or III. To qualify for an upper level the score in each domain must be at least 450. Then, a composite scale score of 600-699 qualifies a teacher for Career Level II and a scale score of 700 or above qualifies one for Level III.

Table 1 lists the fifty plus scores which are calculated in this evaluation system. Scores are first computed in their original metric. Then those scores from the classroom observation, dialogue sessions, student questionnaire, professional skills test, and professional development and leadership summary are translated to standard scores with -3.00 as a lower bound and +3.00 as an upper bound. The weighted standard score mean is calculated for each of the five domains which are then translated to a 200-800 scale (DATA DMN scores) using translation tables derived from standards set at the beginning of the Career Ladder Program. These

scores are then combined with the domain scores from the principal questionnaire and the ratings from the consensus judgment to obtain the domain scores. The composite score is a weighted average of these domain scores. From this composite score teachers are classified into Career Levels I, II, or III.

SAMPLE FOR ANALYSIS

The sample from which data is being analyzed consists of all general education teachers evaluated for the upper levels of the Tennessee Career Ladder during the 1988-89 academic year.

SELECTION OF A METHODOLOGY

There are accepted correlational procedures for determining the degree of internal consistency among test items and scores on a test. However, in this setting, internal consistency among scores and classification in a personnel evaluation system need to be determined. Since there are three Career Levels, it appears some procedure related to analysis of variance would be appropriate. Procedures calculated from ANOVA sums of squares (or mean squares) considered were omega squared (Hays, 1963), the correlation ratio (Glass & Hakstian, 1969), and the intraclass correlation (Haggard, 1958). The appropriate one of these would be most like a test item discrimination statistic. However, for the purposes of the Tennessee Career Ladder Program, it was desirable to determine the amount scores contributed to differentiation between each combination of the three levels. For this reason, effect size as used in the meta-evaluation literature was selected.

ANALYSIS FOR EFFECT SIZE

The calculation of effect size requires the group means and a standard deviation for each score. The standard deviation used for this purpose is the square root of the variance resulting from pooling each group variance. Thus, for a given score, its three measures of effect size are calculated: 1) Career Level II minus Level I, 2) Level III-II, and 3) Level III-I (which is the sum of the first two). Each effect size calculated represents the difference between level means in standard deviation units.

Just as a point biserial correlation for item discrimination is inflated because it is part of the whole score, so these measures of effect size may also be inflated, suggested Linn (1989). Consequently, these effect sizes were calculated a second time using modified Career Ladder Levels. These modified levels were determined by calculating a composite score based on only four of the five domains and then following the same criteria for determined Career Level. Table 2 shows consistency in classification by the actual and modified levels, with a range from 83.8% to 92.9% for correct classifications. Especially important

is that no teacher who was actually assigned Level I would shift to Level III under the modified scoring and also no teacher who was actually assigned Level III would shift to Level I.

RESULTS OF ANALYSIS

Table 1 shows the results of the analysis to determine effect size for each score when the actual Career Ladder Levels are used. This table has a separate line for each of the 66 data scores in the Teacher System of the Tennessee Career Ladder Program. The five DATA DMN scores result from within domain weighted means of the data scores converted to the 200-800 scaled score to be reported to teachers. The five Principal scores are domain means from the original 24 scores calculated from the Principal Questionnaire. The Consensus scores are the four domain scores from the peer evaluator consensus judgement. The Domain scores are the within domain weighted means of the DATA DMN, Principal and Consensus scores. The Composite score is the weighted mean of these five Domain scores. The next column is the proportional weight in the system. The next three columns are the means for those within that Career Ladder Level. The means are followed by the standard deviation calculated from the pooled within group variances. Three columns of effect sizes are then shown. The first is for the comparison of Levels II and I. The second compares Levels III and II. The third is calculated by dividing the difference between the Level III mean and the Level I mean by the standard deviation. Notice that this is also the sum of the first two effect sizes. The last column is for the actual number of teachers with data analyzed on that variable. For most variables the N is 702, but must be less for the scores from the Elementary Student Questionnaire and for the Secondary Student Questionnaire.

The first time this analysis was done, we were surprised at the large values for many of these effect sizes, especially for scores which had weights less than .05 (i.e. 5 percent). Notice that the first score, Observation 101 (Teacher develops higher level thinking skills) has effect sizes over 1.0 when comparing adjacent levels, and thus is 2.20 for comparing Level III and Level I. When comparing Level II to Level I, for the 66 data scores, seventeen effect sizes were above 1.00 and another thirteen were in the range .80 to .99. Only six of these effect sizes were below .30. The effect sizes comparing Levels III and II are not as high, with only eight effect sizes above 1.0 and eight more between .80 to .99. Also twelve were below .30. The only negative effect size was for Observation 125 (Did the teacher show favoritism based on sex or race?), which resulted from 2 of 702 teachers exhibiting this undesirable behavior and both of them still reached Level III. One score had effect sizes of .00 for both comparisons because all teachers received the maximum score on Observation 104 (Use of materials is related to instruction). Even though these two scores do not differentiate among the levels earned, they are maintained

in the system because these behaviors were decided to be important, and those few teachers who did not consistently show the desirable behavior should receive a lower score.

Several interpretive comments reflecting the utility of the effect sizes follow. Some teachers have criticized this system for utilizing scores from student questionnaires, especially from elementary students. However, the effects sizes for comparing Levels II and I range from .51 to 1.43 (with 7 of 9 above 1.00) for the elementary students and from .48 to 1.05 for the secondary students. For differentiating between Levels III and II the effect sizes for elementary students range from .46 to .69 and for secondary students they range from .43 to .77. The principals have also been criticized for not differentiating between teachers at the various levels. With effect size ranging from .66 to .76 for comparing Levels II and I, it appears that they do make this distinction. They do not differentiate as well between Levels III and II with effect sizes of .20 to .36. Examination of the means suggests that the problem with the principal ratings is more a matter of inflated scores than lack of ability to differentiate, at least Level II from Level I. For the principals the Level I means range from 748 to 759 on a system where the Composite score above 700 is Level III and 600-699 is Level II.

The influence of the part to whole relationship was removed by using a modified level in which a teacher's level was determined from a composite score from which the domain for the score under consideration was omitted. Table 4, which is similar to Table 3, shows effect sizes calculated by using this modified classification. Generally this did result in lower effect sizes. When comparing Levels II and I the number of scores with effect sizes above 1.00 decreased from 17 to 15, and the number in the range .80 to .99 decreased from 13 to 10. Also, the number of scores with effect sizes below .30 increased from 6 to 10. When comparing Levels III and II the number of scores with effect sizes above 1.00 decreased from 8 to 7, and the number in the range .80 to .99 decreased from 8 to 2. Also, the number of scores with effect sizes below 30 increased from 12 to 14.

Thus far effect sizes for individual scores have been examined. Another type of information desired dealt with the contribution of each domain to the separation among levels. For this purpose, examine the values for the DOMAIN scores at the end of Table 3. When using the actual Career Ladder Level, for comparing Levels II and I the effect sizes range from 0.84 to 3.18 and for Levels III and II they range from 0.33 to 1.88.

For determination of the total differentiation, the effect sizes for Level III and I range from 1.17 to 4.63. The two largest effects are associated with the domains with the largest weights in the system. Thus, the examination of the DOMAIN effect sizes for the modified Levels are of particular interest, since this

would remove both the part to whole inflation of the relationship and the impact of the larger weight.

Figures 1 - 3 should facilitate this comparison of the impact upon effect size of using actual versus modified levels. Figure 1 shows the comparison of effect sizes for the actual and modified Levels comparing II and I for each DOMAIN score. Two observations can easily be made: 1) the pattern across domains is relatively similar, in that for both sets of histograms in Figure 1 DOMAINS 1 and 5 are low compared to the other three and 2) the magnitude of the effect sizes decreases for all domains. Figure 2 shows the comparison of effect sizes for the comparison of Levels III and II. Once again two observations can be made from the two histograms: 1) the pattern across domains is relatively similar in that DOMAIN 1 is largest and DOMAIN 5 is smallest and 2) the magnitude decreases for all domains. Figure 3 shows the effect sizes for the comparison of Levels III and I, which has patterns more like Figure 1. Also in Figure 3 the observations are 1) DOMAIN 1 has the lowest effect size and the large effect sizes are for DOMAINS 2, 3 and 4 and 2) the magnitude of the effect sizes decreases for all domains.

In conclusion, effect size can be used to indicate the degree to which scores in a complex personnel evaluation system to differentiate among the levels earned. The use of modified level in which that data was excluded from the classification scheme does reduce the magnitude of the effect sizes while generally keeping different scores in approximately the same relative order.

REFERENCES

Glass, G.V & Hakstian, R. (1969). Measures of association in comparative experiments, American Educational Research Journal, 6, 403-414

Haggard, E.A. (1958). Intraclass correlation and the analysis of variance. New York: Dryden.

Hays, W.L. (1963). Statistics for Psychologists. New York: Holt, Rinehart, and Winston.

Linn, R.L. (1989). Personal communication to Tennessee Career Ladder Program staff.

Wolf, F.M. (1986). Meta-analysis: Quantitative Methods for Research Synthesis. Beverly Hills, CA: Sage Publications.

Table 1. Tennessee Career Ladder Domains, Instruments, and Scores

I. DOMAINS OF COMPETENCE

1. PLANNING: Plans and prepares for instruction effectively.
2. TEACHING STRATEGIES: Uses teaching strategies and procedures appropriate to the content, objectives, and learners.
3. EVALUATION: Uses evaluation of student progress to improve instruction.
4. CLASSROOM MANAGEMENT: Manages classroom activities effectively.
5. PROF. DEVELOP. & LEADERSHIP: Establishes and maintains a professional leadership role and improves professional skills and knowledge.
6. COMMUNICATION: Uses communication skills effectively.

II. SOURCES OF DATA

1. Teacher
2. Peer evaluators
3. Prinicpals (or supervisors)
4. Students

III. INSTRUMENTS AND METHODS USE TO COLLECT DATA

1. Observ Classroom observations
2. Dialog Dialogue sessions
3. Studnt Student questionnaire (E=elementary, S=secondary)
4. PST Written test: Tennessee Professional Skills Test
5. PDL Professional development and leadership summary
6. Princip1 Principal questionnaire
7. Consnsus Evaluator consensus

IV. SCORES

Observ 101 Teacher develops higher level thinking skills
Observ 102 Teacher preparation
Observ 103 Uses variety of materials/equipment
Observ 104 Uses of materials is related to instruction
Observ 105 Teacher Behavior: Task and content
Observ 106 Procedural questions asked by students (R)
Observ 107 Did students exhibit lack of understanding? (R)
Observ 108 Did teacher present correct information?
Observ 109 Teacher provides appropriate language model for learners
Observ 110 Teacher Behavior: Monitors
Observ 111 Did teacher respond to lack of understanding by reteaching?
Observ 112 Attends to need in order of importance & paces activities to accomodate learner understanding
Observ 113 Teacher Behavior: Practice and review
Observ 114 Student initiates/responds
Observ 115 Academic questions asked by students
Observ 116 Students remain on task
Observ 117 Teacher involves students who do not volunteer
Observ 118 Student preparation
Observ 119 Teacher encourages student participation
Observ 120 Teacher Behavior: Feedback
Observ 121 Teacher uses feedback and evaluation
Observ 122 Teacher use reasonable sanctions for inappropriate behavior
Observ 123 Teacher control

Observ 124 Positive affect
Observ 125 Did the teacher show favoritism based on sex or race? (R)
Observ 126 Climate: Academic environment conducive to learning & demonstrates appreciation for learner efforts/accomplishments
Observ 127 Time elapsed between interruption and class back on task
Observ 128 Effective use of resources
Observ 129 Effective use of facilities
Observ 130 Teacher Behavior: Other - classroom procedures
Observ 131 Teacher Behavior: Other - teacher off-task (R)

Dialog 201 Planning focus: curriculum
Dialog 202 Planning focus: teaching delivery
Dialog 203 Planning focus: adjusts for individual differences
Dialog 204 Teaching strategy: content
Dialog 205 Teaching strategy: learner differences
Dialog 206 Teaching strategy: practice/review
Dialog 207 Teaching strategy: student involvement
Dialog 208 Evaluation focus: uses assessment for adaptations
Dialog 209 Evaluation focus: records and reporting of student progress
Dialog 210 Evaluation focus: student achievement
Dialog 211 Class management: high expectations and appropriate grouping

EStudnt 501 = SStudnt 601 Teacher explains academic task to learners
EStudnt 502 = SStudnt 602 T. gives clear directions and paces appropriately
EStudnt 503 = SStudnt 603 Provides opportunity to review/practice
EStudnt 504 = SStudnt 604 Reinforces importance of schoolwork
EStudnt 505 = SStudnt 605 T. assists learner through feedback
EStudnt 506 = SStudnt 606 T. promotes student learning/attitude
EStudnt 507 = SStudnt 607 T. promotes appropriate student behavior
EStudnt 508 = SStudnt 608 T. recognizes student effort
EStudnt 509 = SStudnt 609 T. promotes efficient use of time

PST 801 Planning
PST 802 Teaching strategies
PST 803 Evaluation
PST 804 Classroom management

PDL 901 Professional development/improvement
PDL 902 Professional Leadership

Princip1 1 Planning
Princip1 2 Teaching strategies
Princip1 3 Evaluation
Princip1 4 Classroom management
Princip1 5 Professional development and leadership

Consnsus 1 Planning
Consnsus 2 Teaching strategies
Consnsus 3 Evaluation
Consnsus 4 Classroom management

Table 2. Crosstabulation Table with Frequencies of Actual Career Ladder Level and Modified Level

Modified Level: Domains Used	Actual Career Ladder Level I			II			III			Percent Same
	I	II	III	I	II	III	I	II	III	
All five	103	0	0	0	328	0	0	0	271	100.0%
2,3,4,5	86	17	0	2	231	95	0	0	271	83.8%
1,3,4,5	77	26	0	9	288	31	0	38	233	85.2%
1,2,4,5	99	4	0	12	312	4	0	30	241	92.9%
1,2,3,5	94	9	0	21	303	4	0	56	215	87.2%
1,2,3,4	103	0	0	22	306	0	0	35	236	91.8%

TABLE 3. EFFECT SIZE ANALYSIS AMONG ACTUAL CAREER LADDER LEVELS:
1988 - 89 GENERAL EDUCATION TEACHERS

SCORE	WEIGHT	MEAN SCORE FOR LEVEL I	II	III	STANDARD DEVIATION	EFFECT SIZE II-I	III-II	III-I	TOTAL N
N = ABOUT		103	328	271					
Observ 101	.0195	48.33	59.04	69.30	9.55	1.12*	1.08*	2.20*	702
Observ 102	.0090	58.51	69.89	81.02	8.82	1.29*	1.26*	2.55*	702
Observ 103	.0045	20.77	23.43	26.34	7.46	0.36*	0.39*	0.75*	702
Observ 104	.0015	100.00	100.00	100.00	0.00	0.00	0.00	0.00	702
Observ 105	.0128	9.52	10.89	12.32	3.31	0.41*	0.43*	0.85*	702
Observ 106	.0053	81.65	87.32	90.73	14.93	0.38*	0.23*	0.61*	700
Observ 107	.0030	93.17	94.62	95.02	5.14	0.28*	0.08	0.36*	702
Observ 108	.0030	99.56	99.86	99.91	0.91	0.33*	0.06	0.39*	702
Observ 109	.0015	99.36	99.54	99.72	1.89	0.10	0.09	0.19	702
Observ 110	.0128	12.22	14.69	18.60	5.20	0.47*	0.75*	1.23*	702
Observ 111	.0064	98.55	99.78	99.83	1.75	0.70*	0.03	0.73*	702
Observ 112	.0064	50.51	62.00	71.52	7.21	1.59*	1.32*	2.92*	702
Observ 113	.0240	7.32	8.48	9.78	2.75	0.42*	0.48*	0.90*	702
Observ 114	.0075	7.46	8.67	9.88	1.52	0.80*	0.79*	1.59*	702
Observ 115	.0083	58.95	65.54	71.51	21.72	0.30*	0.28*	0.58*	700
Observ 116	.0053	99.95	99.98	99.99	0.04	0.89*	0.27*	1.16*	702
Observ 117	.0038	57.44	66.23	74.47	17.17	0.51*	0.48*	0.99*	702
Observ 118	.0053	50.21	56.34	62.76	7.10	0.86*	0.90*	1.77*	702
Observ 119	.0075	55.73	65.86	75.21	8.87	1.14*	1.05*	2.20*	702
Observ 120	.0188	18.32	22.35	25.23	4.78	0.84*	0.60*	1.45*	702
Observ 121	.0113	50.02	60.45	69.55	7.04	1.48*	1.29*	2.77*	702
Observ 122	.0090	64.13	84.98	97.07	31.05	0.67*	0.39	1.06*	202
Observ 123	.0270	97.49	99.25	99.76	1.52	1.16*	0.34*	1.50*	702
Observ 124	.0120	0.54	0.83	1.16	0.47	0.61*	0.70*	1.31*	702
Observ 125	.0030	100.00	100.00	99.94	0.63	0.00	-0.10	-0.10	702
Observ 126	.0210	54.40	65.99	75.87	7.63	1.52*	1.30*	2.81*	702
Observ 127	.0038	99.89	99.91	99.94	0.11	0.18	0.24	0.43	45
Observ 128	.0075	51.85	60.75	70.56	6.52	1.37*	1.51*	2.87*	702
Observ 129	.0030	52.55	58.31	67.85	8.07	0.71*	1.18*	1.90*	702
Observ 130	.0019	94.50	97.15	98.55	4.71	0.56*	0.30*	0.86*	702
Observ 131	.0019	98.69	99.66	99.88	1.45	0.67*	0.15	0.82*	702
Dialog 201	.0225	2.92	3.22	3.68	0.47	0.65*	0.99*	1.64*	702
Dialog 202	.0165	2.92	3.21	3.65	0.49	0.59*	0.91*	1.50*	702
Dialog 203	.0120	2.81	3.10	3.61	0.59	0.49*	0.87*	1.36*	702
Dialog 204	.0210	3.35	3.65	4.01	0.39	0.77*	0.90*	1.67*	702
Dialog 205	.0210	2.81	3.24	3.69	0.48	0.91*	0.93*	1.83*	702
Dialog 206	.0210	2.96	3.36	3.78	0.48	0.83*	0.89*	1.73*	702
Dialog 207	.0120	2.97	3.40	3.84	0.49	0.88*	0.89*	1.77*	702
Dialog 208	.0150	2.81	3.31	3.66	0.48	1.04*	0.74*	1.78*	702
Dialog 209	.0136	2.80	3.33	3.68	0.54	0.99*	0.64*	1.63*	702
Dialog 210	.0120	2.66	3.17	3.55	0.61	0.84*	0.61*	1.45*	702
Dialog 211	.0075	2.79	3.24	3.62	0.57	0.80*	0.68*	1.48*	702
EStudnt501	.0075	4.44	4.69	4.83	0.24	1.07*	0.58*	1.65*	233
EStudnt502	.0075	4.08	4.42	4.57	0.28	1.18*	0.53*	1.71*	233
EStudnt503	.0090	3.59	3.81	4.07	0.42	0.51	0.63*	1.14*	233
EStudnt504	.0090	3.92	4.25	4.48	0.37	0.91*	0.63*	1.54*	233
EStudnt505	.0075	3.34	3.89	4.15	0.46	1.20*	0.57*	1.77*	233

TABLE 3 (CONT.)

SCORE	WEIGHT	MEAN SCORE FOR LEVEL I	MEAN SCORE FOR LEVEL II	MEAN SCORE FOR LEVEL III	STANDARD DEVIATION	EFFECT SIZE II-I	EFFECT SIZE III-II	EFFECT SIZE III-I	TOTAL N
EStudnt506	.0075	4.20	4.58	4.73	0.32	1.16*	0.46*	1.62*	233
EStudnt507	.0135	4.07	4.45	4.61	0.28	1.36*	0.56*	1.92*	233
EStudnt508	.0165	4.15	4.49	4.65	0.24	1.43*	0.69*	2.12*	233
EStudnt509	.0105	4.10	4.44	4.64	0.30	1.11*	0.68*	1.79*	233
SStudnt601	.0075	4.39	4.62	4.76	0.28	0.82*	0.50*	1.33*	296
SStudnt602	.0075	4.33	4.46	4.57	0.26	0.48*	0.43*	0.91*	296
SStudnt603	.0090	3.65	3.87	4.12	0.37	0.60*	0.69*	1.29*	296
SStudnt604	.0090	3.89	4.09	4.36	0.35	0.56*	0.77*	1.33*	296
SStudnt605	.0075	3.47	3.87	4.16	0.39	1.05*	0.74*	1.79*	296
SStudnt606	.0075	4.08	4.32	4.57	0.37	0.67*	0.68*	1.35*	296
SStudnt607	.0135	3.81	4.02	4.25	0.33	0.61*	0.72*	1.34*	296
SStudnt608	.0165	4.19	4.44	4.63	0.29	0.86*	0.67*	1.53*	296
SStudnt609	.0105	4.46	4.60	4.69	0.21	0.68*	0.43*	1.11*	296
PST 801	.0120	18.76	19.94	20.81	2.86	0.41*	0.31*	0.72*	702
PST 802	.0270	18.75	19.87	20.89	2.43	0.46*	0.42*	0.88*	702
PST 803	.0120	15.32	15.74	16.72	3.25	0.13	0.30*	0.43*	702
PST 804	.0195	20.28	21.37	21.89	2.29	0.47*	0.23*	0.70*	702
PDL 901	.0240	2.83	3.18	3.40	0.58	0.59*	0.37*	0.96*	702
PDL 902	.0560	2.59	2.93	3.18	0.59	0.58*	0.43*	1.01*	702
Consnsus 1	.0375	578.83	673.62	730.42	52.87	1.79*	1.07*	2.87*	702
Consnsus 2	.0875	565.42	673.84	732.94	56.32	1.93*	1.05*	2.97*	702
Consnsus 3	.0375	561.32	663.92	724.20	58.41	1.76*	1.03*	2.79*	702
Consnsus 4	.0625	562.15	685.70	742.72	71.82	1.72*	0.79*	2.51*	702
Principl 1	.0150	755.93	778.96	791.41	35.04	0.66*	0.36*	1.01*	702
Principl 2	.0150	753.87	779.75	791.36	35.89	0.72*	0.32*	1.04*	702
Principl 3	.0150	746.80	776.03	786.65	37.79	0.77*	0.28*	1.05*	702
Principl 4	.0300	753.84	781.19	791.19	37.24	0.73*	0.27*	1.00*	702
Principl 5	.0200	746.96	776.03	783.53	38.22	0.76*	0.20	0.96*	702
Data DMN 1	.0975	424.01	494.38	636.13	80.74	0.87*	1.76*	2.63*	702
Data DMN 2	.2475	494.23	633.81	708.28	52.18	2.68*	1.43*	4.10*	702
Data DMN 3	.0975	523.68	675.41	722.42	59.07	2.57*	0.80*	3.36*	702
Data DMN 4	.1575	507.21	680.77	727.60	56.48	3.07*	0.83*	3.90*	702
Data DMN 5	.0800	697.10	726.44	738.75	41.41	0.71*	0.30*	1.01*	702
DOMAIN 1	.1500	495.91	567.64	675.23	57.16	1.26*	1.88*	3.14*	702
DOMAIN 2	.3500	522.41	649.65	717.77	42.23	3.01*	1.61*	4.63*	702
DOMAIN 3	.1500	555.40	682.60	729.29	46.22	2.75*	1.01*	3.76*	702
DOMAIN 4	.2500	550.54	694.05	739.01	45.09	3.18*	1.00*	4.18*	702
DOMAIN 5	.1000	707.07	736.36	747.70	34.86	0.84*	0.33*	1.17*	702
COMPOSITE	1.0000	548.82	662.07	721.46	26.93	4.21*	2.21*	6.41*	702

LEVEL I: COMPOSITE LESS THAN 600; OR AT LEAST ONE DOMAIN LESS THAN 450
LEVEL II: COMPOSITE GREATER THAN OR EQUAL TO 600 AND LESS THAN 700; AND ALL DOMAIN SCORES 450 OR ABOVE
LEVEL III: COMPOSITE GREATER THAN OR EQUAL TO 700; AND ALL DOMAIN SCORES 450 OR ABOVE

TABLE 4. EFFECT SIZE ANALYSIS AMONG MODIFIED CAREER LADDER LEVELS: 1988 - 89 GENERAL EDUCATION TEACHERS

SCORE	DOMAIN	MEAN SCORE FOR LEVEL I	II	III	STANDARD DEVIATION	EFFECT SIZE II-I	III-II	III-I	TOTAL N
N = ABOUT		103	328	271					
Observ 101	1	47.34	57.65	67.38	9.73	1.06*	1.00*	2.06*	702
Observ 102	1	57.73	68.24	78.97	9.07	1.16*	1.18*	2.34*	702
Observ 103	1	21.38	23.08	25.57	7.55	0.23	0.33*	0.56*	702
Observ 104	1	100.00	100.00	100.00	0.00	0.00	0.00	0.00	702
Observ 105	2	9.84	10.94	12.10	3.37	0.33*	0.34*	0.67*	702
Observ 106	2	80.41	87.37	90.78	14.89	0.47*	0.23*	0.70*	700
Observ 107	2	93.24	94.35	95.29	5.13	0.22	0.18	0.40*	702
Observ 108	2	99.54	99.87	99.90	0.91	0.35*	0.03	0.39*	702
Observ 109	2	99.33	99.50	99.77	1.89	0.09	0.14	0.23	702
Observ 110	2	12.40	14.90	18.21	5.34	0.47*	0.62*	1.09*	702
Observ 111	2	98.49	99.75	99.81	1.75	0.72*	0.03	0.75*	702
Observ 112	2	49.59	62.02	71.30	7.41	1.68*	1.25*	2.93*	702
Observ 113	2	7.39	8.60	9.55	2.80	0.43*	0.34*	0.77*	702
Observ 114	2	7.56	8.76	9.68	1.60	0.75*	0.57*	1.33*	702
Observ 115	2	59.87	65.44	71.08	21.82	0.26	0.26*	0.51*	700
Observ 116	2	99.93	99.98	99.99	0.04	1.23*	0.30*	1.53*	702
Observ 117	2	57.52	66.33	73.96	17.34	0.51*	0.44*	0.95*	702
Observ 118	2	49.76	56.43	62.56	7.21	0.92*	0.85*	1.77*	702
Observ 119	2	54.05	66.34	74.71	8.99	1.37*	0.93*	2.30*	702
Observ 120	3	18.91	22.49	25.20	4.87	0.73*	0.56*	1.29*	702
Observ 121	3	50.67	61.16	69.56	7.30	1.44*	1.15*	2.59*	702
Observ 122	4	68.99	85.24	93.44	32.13	0.51*	0.26	0.76*	202
Observ 123	4	97.80	99.28	99.77	1.56	0.95*	0.31*	1.26*	702
Observ 124	4	0.57	0.87	1.18	0.48	0.62*	0.66*	1.28*	702
Observ 125	4	100.00	100.00	99.92	0.63	0.00	-0.12	-0.12	702
Observ 126	4	55.64	67.09	76.35	8.00	1.43*	1.16*	2.59*	702
Observ 127	4	99.90	99.91	99.95	0.11	0.10	0.34	0.44	45
Observ 128	4	52.25	61.72	71.54	6.53	1.45*	1.50*	2.95*	702
Observ 129	4	53.00	59.31	68.52	8.22	0.77*	1.12*	1.89*	702
Observ 130	4	94.83	97.29	98.61	4.73	0.52*	0.28*	0.80*	702
Observ 131	4	98.95	99.61	99.93	1.47	0.45*	0.21*	0.67*	702
Dialog 201	1	2.94	3.21	3.56	0.49	0.53*	0.71*	1.24*	702
Dialog 202	1	2.95	3.22	3.51	0.52	0.52*	0.56*	1.07*	702
Dialog 203	1	2.86	3.06	3.48	0.61	0.33*	0.68*	1.00*	702
Dialog 204	2	3.44	3.66	3.96	0.42	0.53*	0.72*	1.25*	702
Dialog 205	2	2.93	3.25	3.63	0.51	0.62*	0.75*	1.37*	702
Dialog 206	2	3.07	3.39	3.69	0.52	0.62*	0.60*	1.21*	702
Dialog 207	2	3.09	3.42	3.75	0.53	0.62*	0.64*	1.26*	702
Dialog 208	3	2.89	3.33	3.65	0.50	0.90*	0.63*	1.53*	702
Dialog 209	3	2.90	3.34	3.67	0.56	0.79*	0.59*	1.38*	702
Dialog 210	3	2.73	3.18	3.55	0.62	0.73*	0.59*	1.32*	702
Dialog 211	4	2.81	3.28	3.65	0.57	0.82*	0.65*	1.47*	702
EStudnt501	2	4.40	4.69	4.82	0.24	1.22*	0.55*	1.77*	233
EStudnt502	2	4.02	4.42	4.56	0.28	1.44*	0.48*	1.93*	233
EStudnt503	2	3.46	3.82	4.06	0.41	0.89*	0.58*	1.47*	233
EStudnt504	2	3.80	4.26	4.47	0.36	1.27*	0.60*	1.87*	233
EStudnt505	3	3.34	3.91	4.17	0.46	1.25*	0.56*	1.81*	233

TABLE 4 (CONT.)

SCORE	DOMAIN	MEAN SCORE FOR LEVEL I	MEAN SCORE FOR LEVEL II	MEAN SCORE FOR LEVEL III	STANDARD DEVIATION	EFFECT SIZE II-I	EFFECT SIZE III-II	EFFECT SIZE III-I	TOTAL N
EStudnt506	3	4.23	4.58	4.74	0.32	1.08*	0.52*	1.60*	233
EStudnt507	4	4.18	4.46	4.60	0.30	0.94*	0.47*	1.41*	233
EStudnt508	4	4.24	4.50	4.66	0.25	1.04*	0.60*	1.64*	233
EStudnt509	4	4.17	4.47	4.64	0.31	0.97*	0.55*	1.52*	233
SStudnt601	2	4.36	4.63	4.76	0.27	0.98*	0.49*	1.47*	296
SStudnt602	2	4.30	4.47	4.58	0.26	0.65*	0.42*	1.07*	296
SStudnt603	2	3.65	3.89	4.12	0.37	0.66*	0.63*	1.29*	296
SStudnt604	2	3.81	4.14	4.36	0.35	0.93*	0.64*	1.57*	296
SStudnt605	3	3.44	3.93	4.14	0.39	1.25*	0.52*	1.78*	296
SStudnt606	3	4.06	4.36	4.57	0.37	0.83*	0.56*	1.39*	296
SStudnt607	4	3.88	4.06	4.21	0.35	0.52*	0.43*	0.95*	296
SStudnt608	4	4.26	4.46	4.61	0.31	0.67*	0.47*	1.14*	296
SStudnt609	4	4.48	4.62	4.67	0.21	0.67*	0.20	0.87*	296
PST 801	1	19.13	19.79	20.55	2.90	0.23	0.26*	0.49*	702
PST 802	2	19.23	19.80	20.77	2.48	0.23	0.39*	0.62*	702
PST 803	3	15.50	15.82	16.64	3.26	0.10	0.25*	0.35*	702
PST 804	4	20.63	21.32	21.97	2.31	0.30*	0.28*	0.58*	702
PDL 901	5	2.90	3.21	3.39	0.59	0.51*	0.30*	0.82*	702
PDL 902	5	2.70	2.95	3.18	0.60	0.42*	0.39*	0.81*	702
Consnsus 1	1	570.44	661.41	722.08	52.41	1.74*	1.16*	2.89*	702
Consnsus 2	2	564.44	672.86	729.16	60.24	1.80*	0.93*	2.73*	702
Consnsus 3	3	570.38	668.18	723.83	60.95	1.60*	0.91*	2.52*	702
Consnsus 4	4	576.48	693.24	742.84	75.18	1.55*	0.66*	2.21*	702
Principl 1	1	756.02	777.11	788.47	35.37	0.60*	0.32*	0.92*	702
Principl 2	2	750.50	779.23	791.79	35.76	0.80*	0.35*	1.15*	702
Principl 3	3	748.18	777.33	786.27	37.91	0.77*	0.24*	1.00*	702
Principl 4	4	759.75	781.78	790.97	37.82	0.58*	0.24*	0.83*	702
Principl 5	5	751.40	777.06	783.51	38.46	0.67*	0.17	0.83*	702
Data DMN 1	1	430.06	489.39	598.38	94.05	0.63*	1.16*	1.79*	702
Data DMN 2	2	515.10	630.41	699.00	66.33	1.74*	1.03*	2.77*	702
Data DMN 3	3	543.08	676.98	721.37	64.94	2.06*	0.68*	2.75*	702
Data DMN 4	4	543.97	682.26	726.42	68.89	2.01*	0.64*	2.65*	702
Data DMN 5	5	704.09	727.23	738.47	41.98	0.55*	0.27*	0.82*	702
DOMAIN 1	1	497.75	561.16	648.31	66.90	0.95*	1.30*	2.25*	702
DOMAIN 2	2	536.85	646.98	710.25	54.88	2.01*	1.15*	3.16*	702
DOMAIN 3	3	570.42	684.81	728.47	51.26	2.23*	0.85*	3.08*	702
DOMAIN 4	4	577.99	696.95	738.28	54.94	2.17*	0.75*	2.92*	702
DOMAIN 5	5	713.55	737.19	747.48	35.44	0.67*	0.29*	0.96*	702
COMPOSITE		548.82	662.07	721.46	26.93	4.21*	2.21*	6.41*	702

LEVEL I: COMPOSITE LESS THAN 600; OR AT LEAST ONE DOMAIN LESS THAN 450
LEVEL II: COMPOSITE GREATER THAN OR EQUAL TO 600 AND LESS THAN 700; AND ALL DOMAIN SCORES 450 OR ABOVE
LEVEL III: COMPOSITE GREATER THAN OR EQUAL TO 700; AND ALL DOMAIN SCORES 450 OR ABOVE

Figure 1

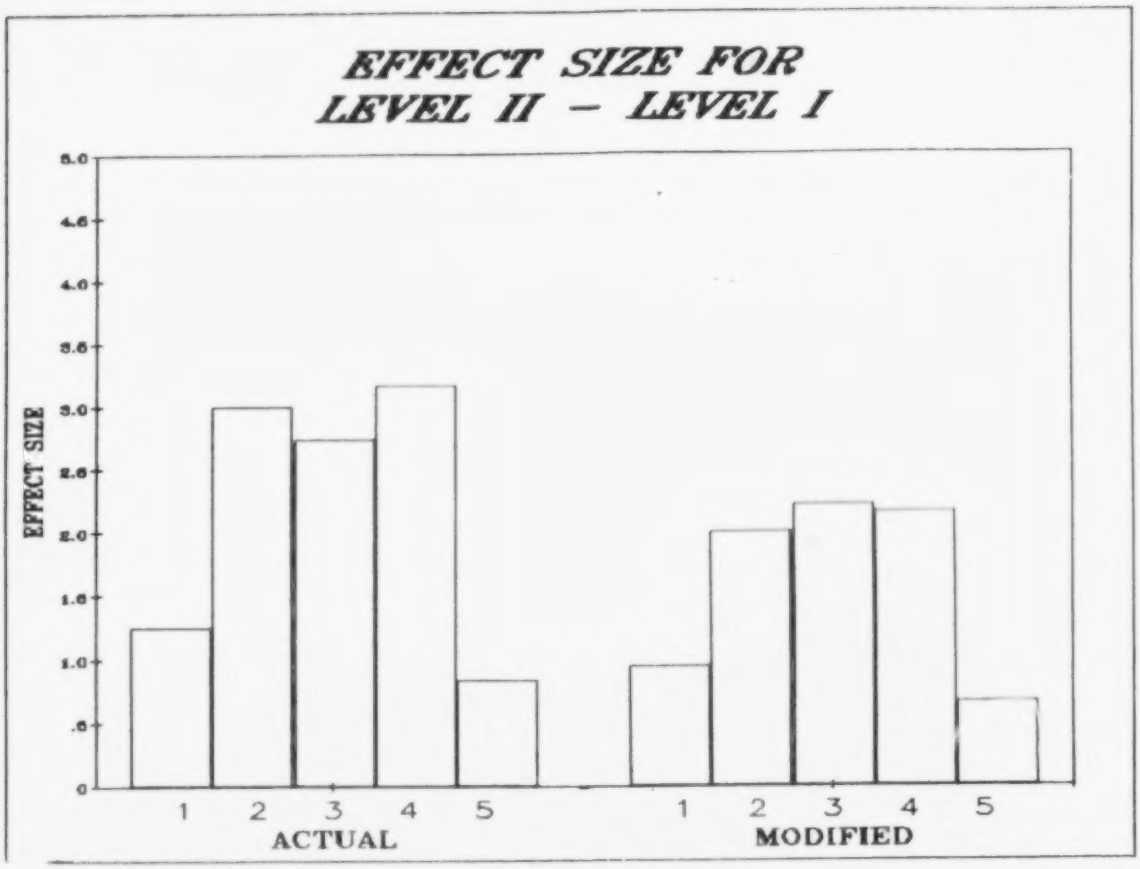


Figure 2

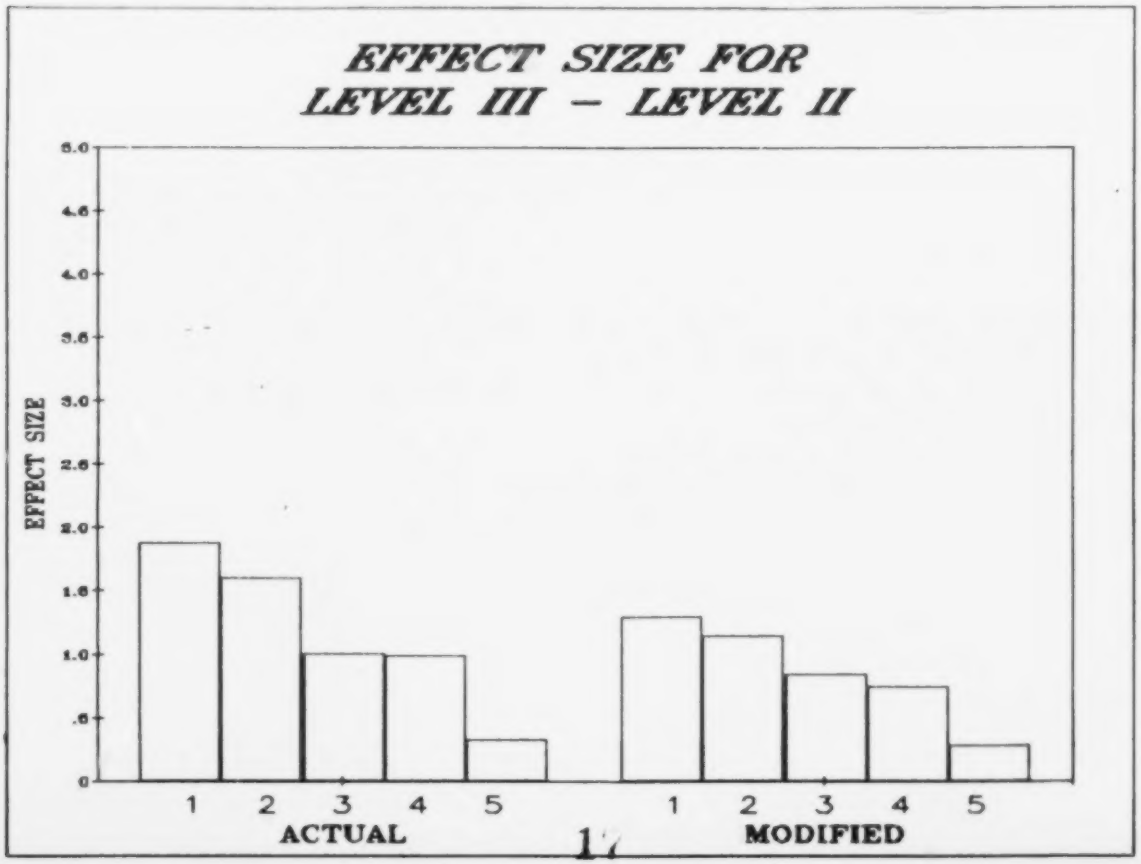
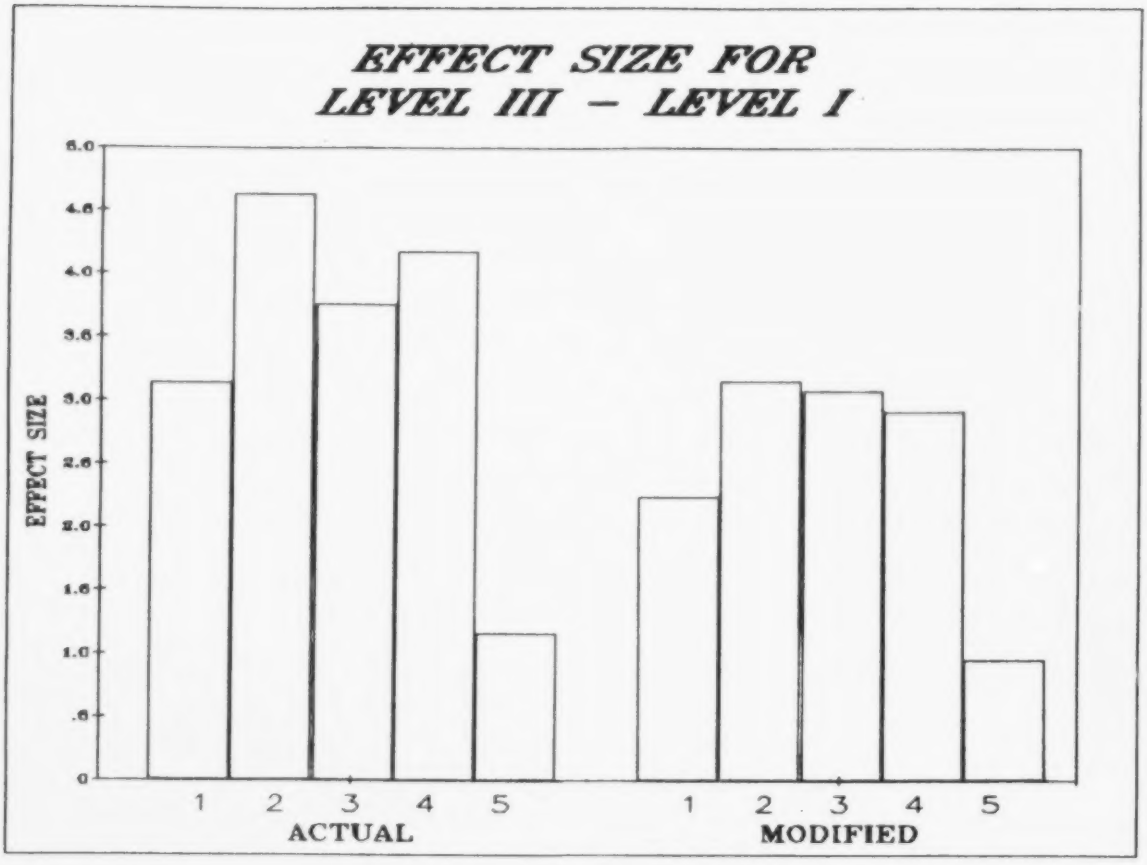


Figure 3


EFFECT SIZE FOR
LEVEL III – LEVEL I
EFFECT SIZE
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
.5
0
1 2 3 4 5
ACTUAL
1 2 3 4 5
MODIFIED